IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

7/Brief
5-15-03

I hereby certify that this correspondence is being deposited with the United States Postal Service as FIRST CLASS MAIL in an envelope addressed to: Commissioner of Patents and Trademarks, Washington, DC 20231 on:

4-7-03
Date of Mailing

Carol Ann Mahoney
Signature

APPEAL BRIEF OF EMRE GUNDUZHAN
FOR
APPARATUS AND METHOD FOR REGENERATING A LOST AUDIO SIGNAL

Serial No. 09/353,906
Filed: July 15, 1999

Appeal from a decision of the Primary Examiner dated September 23, 2002
Technology Center 2654
Examiner M. Opsasnick

TABLE OF CONTENTS

TABLE OF AUTHORITIES

I. Real Party in Interest

The real party in interest is Nortel Networks, Limited.

II. Related Appeals and Interferences

Appellant is not aware of any appeals or interferences that are related to the present case.

III. Status of the Claims

This is an appeal brief from a decision by the Primary Examiner dated September 23, 2002, finally rejecting claims 1 - 31, currently pending in the present application. No claims have been allowed. Claims 1-31 are the subject of this appeal.

A notice of Appeal was filed on January 23, 2003.

IV. Status of Amendments

In an Office Action dated April 16, 2002, all pending claims 1 – 31 were rejected under 35 U.S.C. §103. On July 12, 2002, Appellant filed a response pursuant to 37 C.F.R.§1.111 which did not amend any of the claims. In the Final Office Action dated September 23, 2002, the rejection of claims 1-31 under 35 U.S.C.§103 was maintained. A Notice of Appeal was filed on January 23, 2003.

V. Summary of the Invention

A. Background

The invention generally relates to data transmission networks and more particularly to regenerating an audio signal segment in an audio signal transmitted across a data transmission network.

Network devices on the Internet commonly transmit audio signals to other network devices ("receivers") on the Internet. To that end, prior to transmission, a given audio signal commonly is divided into a series of contiguous audio segments that each are encapsulated within one or more Internet Protocol packets. Each segment includes a plurality of samples that identify the amplitude of the signal at specific times. Once filled with one or more audio segments, each Internet Protocol packet is transmitted to one or more Internet receiver(s) in accord with the well known Internet Protocol.

As known in the art, Internet Protocol packets commonly are lost during transmission across the Internet. Undesirably, the loss of Internet Protocol packets transporting audio segments often significantly degrades signal quality to unacceptable levels. This problem is further exasperated when transmitting a real-time voice signal across the Internet, such as a real-time voice signal transmitted during a teleconference conducted across the Internet.

B. Appellant's Invention

Appellant provides a mechanism for generating a new audio segment that is based upon a given lost audio segment ("given segment") of an audio signal. This mechanism advantageously improves voice quality in an environment where packets carrying voice information may be lost. In an environment in which the Appellant's invention may be employed, voice information is

digitized and broken down into segments which are transported across a network encapsulated for example in Internet Protocol ("IP") packets. Some segments are sometimes lost. Segment generators in apparatus that receive the packets utilize previously received audio segments to regenerate approximations of lost audio segments of a received audio signal.

For example, a first telephone may receive a plurality of Internet Protocol packets ("IP packets") transporting a given real-time voice signal from a second telephone. Upon analysis of the received IP packets, the first telephone may detect that it had not received all of the necessary IP packets to reproduce the entire given signal. Such IP packets that were not received may have been lost during transmission, thus losing one or more audio segments of the given audio (voice) signal. As detailed below, a segment generator in the first telephone regenerates the missing one or more audio segments from the received audio segments to produce a set of regenerated audio segments. The set of regenerated audio segments, however, is an approximation of the lost audio segments and thus, is not necessarily an exact copy of such segments. Once generated, each segment in the set of regenerated audio segments is added to the given audio signal in its appropriate location, thus reconstructing the entire signal. If subsequent audio segments are similarly lost, the regenerated segment can be utilized to regenerate such subsequent audio segments.

More particularly, the segment generator receives previous segments of the audio signal. A linear predictive coding analyzer ("LP analyzer") within the segment generator determines the characteristics of the formant of the received segments. The LP analyzer forwards the determined formant characteristics to a linear predictive filter ("LPC filter") that utilizes such characteristics to remove the formant from the input segments. The LP analyzer also forwards

the determined formant characteristics to an inverse linear predictive filter ("inverse LPC filter") that restores the formant characteristics to a residue signal (a/k/a "residue segment(s)").

The segment generator also includes a pitch detector that determines the pitch of one or more residue segments, and an estimator that utilizes the determined pitch and residue segments to estimate the residue segments of the lost audio segments being regenerated. An overlap-add module/scaling module are also included to perform conventional overlap-add operations, and conventional scaling operations.

The above elements are used by the segment generator in a process for regenerating the lost audio segment(s) of a real-time voice signal. This process makes use of the symmetric nature of a person's vocal tract over a relatively short time interval. More particularly, according to many well known conventions, a final voice signal is modeled as being a waveform traversing through a tube. The tube is a person's vocal tract, which includes the throat and mouth. When passing through the vocal tract, the waveform is modified by the resonances of the tract, thus producing the final voice signal. The effect of the vocal tract on the waveform thus is represented by the resonances that it produces.. These resonances are known in the art as "formants." Accordingly, removing the formant from a final voice signal produces the original waveform, which is known in the art as a "residue" or a "residue signal." The residue signal may be referred to herein as a set of residue segments.

An audio signal is broken into a sequence of consecutive audio segments for transmission across an IP network. The process used by the segment generator is therefore initiated when it is detected that one of the audio segments is missing from the received sequence of consecutive audio segments. The process therefore begins by retrieving a set of consecutive audio segments that precede the lost segment. The set of retrieved audio segments preferably ranges from one

audio segment to fifteen audio segments, though the segment generator may be preconfigured to utilize any set number of audio segments.

Once the set of audio segments is retrieved, the LP analyzer calculates the tract data (*i.e.*, formant data) from the set of segments, and forwards such data to the LPC filter and inverse LPC filter. The formants are then removed from the input set of audio segments. To that end, the set of audio segments are filtered by the LPC filter produce a set of residue segments. The set of residue segments then are forwarded to both the estimator and pitch detector. The pitch period of the set of residue segments is determined by the pitch detector and forwarded to the estimator. Once received by the estimator, both the determined pitch period and the set of residue segments are processed to produce a new set of residue segments (a/k/a "residue signal") that approximate both a set of residue segments of the lost audio segments, and the residues of the two overlap segments that immediately precede and follow the lost audio segment.

Once the estimator generates the residue of the lost segments, the vocal tract data is added back into the newly generated set of residue segments. To that end, the newly generated set of residue segments is passed through the inverse LPC filter, thus adding the formants of the initially calculated vocal tract. This produces a reproduced set of audio segments that approximate the lost set of audio segments. Once the reproduced set of audio segments is generated, it immediately may be added to the audio signal, thus providing an approximation of the entire audio signal.

VI. Issue

A. Whether claims 1-31 were properly rejected under 35 U.S.C. § 103, over Yeldener, U.S. Patent No. 5,890,108, in view of Shoham, U.S. Patent No. 5,699,485.

VII. Grouping of Claims

Claims 1 - 31 do not stand or fall together. Claims 1 - 31 will be argued separately.

VIII. Argument

A. The Examiner has failed to establish a prima facie case of obviousness under 35 U.S.C. §103 of claims 1-31 as being unpatentable over Yeldener in view of Shoham.

"To establish a *prima facie* case of obviousness, three basic criteria must be met. First, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Second, there must be a reasonable expectation of success. Finally, the prior art reference (or references when combined) must teach or suggest all of the claim limitations." (M.P.E.P. §2143: Basic Requirements of Patentability)

i. Claims 1-10:

1. There is no motivation for the combination of the references

It is well established that, in order to support a rejection under 35 U.S.C. §103, sufficient motivation for combining the references must be shown by the Examiner. The mere fact that references can be combined or modified does not render the resultant combination obvious unless the prior art also suggests the desirability of the combination. *In re Mills*, 916 F.2d 680, 16 USPQ2d 1430 (Fed. Cir. 1990). In determining the propriety of the Patent Office case for

obviousness in the first instance, it is necessary to ascertain whether or not the reference teachings would appear to be sufficient for one of ordinary skill in the relevant art having the reference before him to make the proposed substitution, combination, or other modification. *In re Linter*, 458 F.2d 1013, 1016, 173 USPQ 560, 568 (CCPA 1972).

As the Office Action points out, Yeldener fails to teach or suggest a method for producing a new audio segment including a step of determining that a given audio segment is unascertainable. The Office Action suggests however that Shoham teaches tracking lost frames and using previous information to regenerate such information, and that it would be obvious to "modify the teachings of Yeldener with the lost frame tracking and recovery techniques as taught by Shoham *because it would advantageously improve the reliability of the recovered speech information*." The Applicant respectfully disagrees.

1.a.) <u>Yeldener does not suggest any motivation for modification or combination with Shoham</u>

Yeldener is directed towards the provision of a speech encoding technique that provides high quality voice reconstruction at low to very low bit rates on the basis of a voicing probability determination. (Yeldener Col. 3 lines 59-62.) In Yeldener, it is explained that transitions that fall within a single frame cannot be represented accurately (col. 19 lines 65-67). It is explained in Yeldener that:

> "...one approach to satisfying this tradeoff is the use of frame-to-frame LPC interpolation. Generally, the idea is to achieve an improved spectrum representation by evaluating intermediate sets of parameters between frames, so that transitions are introduced more smoothly at the frame edges without the need to increase the coding capacity." (col. 20 lines 18-23)

Yeldener makes use of information in neighboring frames in addition to that contained in a current frame to improve spectrum representation. In Yeldener, all frames must be present in order for the interpolation techniques to be effective. There is no suggestion in Yeldener that any of the frames are unascertainable – e.g. missing, corrupted, etc., and in fact Yeldener does not address the issues associated with unascertainable segments or frames. Because Yeldener teaches the benefits of smoothing frame edges by using interpolation, Appellant can glean no motivation from Yeldener for the modification necessitated by the Examiner to meet the limitation of all independent claims of "determining that a given audio signal is unascertainable…" using the steps outlined in claim 1 and the other independent claims.

1.b.) <u>Shoham does not suggest any motivation for combination with Yeldener</u>

Similarly, Shoham provides no motivation for the modification suggested by the Examiner. Shoham is directed to reconstructing codebook gain information. The Office Action points to column 2 lines 50 – 60 of Shoham in support of the motivation to modify the references. Part of that portion of Shoham recites as follows:

> "The present invention addresses the problem of the lack of codebook gain information during frame erasure. In accordance with the present invention, a codebook-based speech decoder which fails to receive reliably at least a portion of the current frame of compressed speech information uses a codebook gain which is an attenuated version of a gain from a previous frame of speech. An illustrative embodiment of the present invention is a speech decoder which includes a codebook memory and a signal amplifier. The memory and amplifier are use in (sic) generating a decoded speech signal based on compressed speech information."

Shoham teaches that previously received codebook gain information can be used to produce an attenuated version of codebook gain information, but the Appellant notes that this

method is not extended to data reconstruction. With regard to the lost *data* , Shoham explicitly describes at column 6 that the data is reconstructed using either adaptive or fixed codebooks. Accordingly, no motivation can be found in Shoham, as stated above with regard to Yeldener, for the modification of the claims as suggested by the Examiner. Further, nowhere does Shoham suggest that this teaching should be combined with a system such as that of Yeldener, wherein 1) no information is missing and 2) codebook gain information is not used.

1.c.) <u>The combination as a whole makes no suggestion for combination.</u>

Neither Yeldener nor Shoham provide any motivation for their combination. Furthermore, the deliberate combination of the two references still fails to suggest or describe Appellant's claimed invention of "… *generating a new audio segment for an audio signal*, the audio signal having a plurality of audio segments, the method comprising the steps of determining that a given audio segment is not ascertainable, the location of the given audio segment within the audio signal being ascertainable; locating a set of consecutive audio segments in the audio signal, the set of consecutive audio segments preceding the given audio segment and having a formant; removing the formant from the set of audio segments to produce a set of residue segments having a pitch; processing the pitch and the set of residue segments to produce a new set of residue segments; and adding the formant of the consecutive set of audio segments to the new set of residue segments to produce an output audio segment…" As already stated above, Yeldener teaches a method of generating an audio signal that uses interpolation techniques with the current frame and neighboring frames. Shoham teaches a method of generating an audio signal using fixed or adaptive codebooks. Thus, the combination does not teach the basic limitations of the claim, and therefore the rejection should be withdrawn.

2. <u>The Examiner has failed to provide a convincing line of reasoning as to why the claimed invention is obvious.</u>

To support the conclusion that the claimed invention is directed to obvious subject matter, either the references must expressly or impliedly suggest the claimed invention or the examiner must present a convincing line of reasoning as to why the artisan would have found the claimed invention to have been obvious in light of the teachings of the references. *Ex parte Clapp*, 227 USPQ 972, 973 (Bd. Pat. App. & Inter. 1985).

The Examiner states that it would be obvious to "modify the teachings of Yeldener with the lost frame tracking and recovery techniques as taught by Shoham because it would advantageously improve the reliability of the recovered speech information." The Examiner's suggested motivation is not understood. The Appellant is unclear as to what is meant by "improve the reliability of the recovered speech information". Yeldener has no "recovered speech information" because Yeldener has no unascertainable frames to recover. In fact, Yeldener teaches the benefits of using the existing frame data. Shoham already teaches a method for providing recovered data (from the fixed or adaptive codebooks), and it is not seen how any modification as suggested by the Examiner would improve the reliability of Shoham. Therefore, the Appellant asserts that the Examiner has failed to present a convincing line of reasoning as to why the artisan would have found the claimed invention to have been obvious in light of the teachings of Yeldener and Shoham.

3. <u>The combination presents no reasonable expectation of success</u>

The teaching or suggestion to make the claimed combination and the reasonable expectation of success must be found in the prior art, and not based on applicant's disclosure. *In re Vaeck,* 947 F.2d 488, 20 USPQ2d 1438 (Fed. Cir. 1991).

The Examiner states that it would be obvious to "modify the teachings of Yeldener with the lost frame tracking and recovery techniques as taught by Shoham because it would advantageously improve the reliability of the recovered speech information." The Appellant asserts that the suggested combination has no reasonable expectation of successfully improving the reliability of the recovered speech information, and therefore certainly has no reasonable expectation of success towards suggesting the Appellant's claimed invention. The Appellant asserts that one skilled in the art could successfully combine Yeldener with Shoham. Shoham teaches a method of providing substitute excitation signals during frame erasure for a standard CELP encoder, wherein each frame is determined to be voiced or unvoiced. According to Shoham, "the generation of a substitute excitation signal during periods of frame erasure is dependent on whether the erased frame is classified as voiced (periodic) or unvoiced (aperiodic)." (Shoham Col. 6 lines 2-5.) According to the teachings of Yeldener, each frame is associated with a voicing probability, and voiced and unvoiced portions of each frame are dealt with separately (Col. 3 lines 63-70). Thus, the frames of Shoham are incompatible with the frames of Yeldener, so any method suggested in Shoham for reconstruction of frames cannot be successfully applied to Yeldener.

Accordingly, because the Appellant can find no clear basis for combining the references as suggested by the Examiner, the rejection under 35 U.S.C. 103 for claim 1 and its dependent claims 2 – 9 should be withdrawn.

ii. Claims 11 – 20 are Computer Program Product claims that contain parallel limitations to the Method claims 1 – 10. Accordingly, for reasons similar to those set forth above with regard to claims 1 – 10, Claims 11 – 20 are patentably distinct over Yeldener, Shoham, and any combination thereof, and the rejection should be withdrawn.

iii. Claims 21 – 31 are Apparatus claims that contain parallel limitations to the Method claims 1 – 10. Accordingly, for reasons similar to those set forth above with regard to claims 1 – 10, Claims 11 – 20 are patentably distinct over Yeldener, Shoham, and any combination thereof, and the rejection should be withdrawn.

IX. Conclusion

Appellant submits therefore that the rejection of claims 1-31 under 35 U.S.C. § 103 is improper for failing to provide sufficient motivation to combine the two references. It is therefore respectfully requested that the Board reverse the Examiner's rejections under 35 U.S.C. §103.

Respectfully submitted,

NORTEL NETWORKS LTD.

By: Mary Steubing
Mary Steubing
Reg. No. 37,946
Attorney for Assignee

Date: April 7, 2003

Steubing McGuinness & Manaras LLP
30 Nagog Park
Acton MA 01720
(978) 264-6664

APPENDIX A

1. A method of generating a new audio segment for an audio signal, the audio signal having a plurality of audio segments, the method comprising:

determining that a given audio segment is not ascertainable, the location of the given audio segment within the audio signal being ascertainable;

locating a set of consecutive audio segments in the audio signal, the set of consecutive audio segments preceding the given audio segment and having a formant;

removing the formant from the set of audio segments to produce a set of residue segments having a pitch;

processing the pitch and the set of residue segments to produce a new set of residue segments; and

adding the formant of the consecutive set of audio segments to the new set of residue segments to produce an output audio segment.

2. The method as defined by claim 1 wherein the given audio segment is missing from the plurality of audio segments.

3. The method as defined by claim 1 wherein the audio signal is a voice signal transmitted across a packet based network, the audio signal being a stream of data packets.

4. The method as defined by claim 1 further comprising:

determining the pitch of the set of residue segments.

5. The method as defined by claim 1 wherein the formant is removed by utilizing linear predictive coding filtering techniques.

6. The method as defined by claim 1 wherein the pitch and set of residue segments are processed by utilizing linear predictive coding filtering techniques.

7. The method as defined by claim 1 wherein the formant is a function having a variable value across the set of audio segments.

8. The method as defined by claim 1 further comprising:

applying overlap-add operations to the output audio segment to produce an overlap audio segment.

9. The method as defined by claim 8 further comprising:

scaling the overlap audio segment to produce a scaled audio segment, the scaled audio segment being the new audio segment.

10. The method as defined by claim 1 further comprising:

adding the output audio segment to the audio signal in place of the given audio segment.

11. A computer program product for use on a computer system for generating a new audio segment for an audio signal, the audio signal having a plurality of audio segments, the computer program product comprising a computer usable medium having computer readable program code thereon, the computer readable program code including:

program code for determining that a given audio segment is not ascertainable, the location of the given audio segment within the audio signal being ascertainable;

program code for locating a set of consecutive audio segments in the audio signal, the set of consecutive audio segments preceding the given audio segment and having a formant;

program code for removing the formant from the set of audio segments to produce a set of residue segments having a pitch;

program code for processing the pitch and the set of residue segments to produce a new set of residue segments; and

program code for adding the formant of the consecutive set of audio segments to the new set of residue segments to produce an output audio segment.

12. The computer program product as defined by claim 11 wherein the given audio segment is missing from the plurality of audio segments.

13. The computer program product as defined by claim 11 wherein the audio signal is a voice signal transmitted across a packet based network, the audio signal being a stream of data packets.

14. The computer program product as defined by claim 11 further comprising:

program code for determining the pitch of the set of residue segments.

15. The computer program product as defined by claim 11 wherein the program code for removing the formant comprises program code for utilizing linear predictive coding filtering

techniques.

16. The computer program product as defined by claim 11 wherein the program code for processing includes program code for utilizing linear predictive coding filtering techniques.

17. The computer program product as defined by claim 11 wherein the formant is a function having a variable value across the set of audio segments.

18. The computer program product as defined by claim 11 further comprising:

program code for applying overlap-add operations to the output audio segment to produce an overlap audio segment.

19. The computer program product as defined by claim 18 further comprising:

program code for scaling the overlap audio segment to produce a scaled audio segment, the scaled audio segment being the new audio segment.

20. The computer program product as defined by claim 11 further comprising:

program code for adding the output audio segment to the audio signal in place of the given audio segment.

21. An apparatus for generating a new audio segment for an audio signal, the audio signal having a plurality of audio segments, the apparatus comprising:

a detector for determining that a given audio segment is not ascertainable, the location of the given audio segment within the audio signal being ascertainable;

an input to receive a set of consecutive audio segments, the set of consecutive audio segments preceding the given audio segment;

a filter operatively coupled with the input, the filter removing the formant from the set of consecutive audio segments to produce a set of residue segments having a pitch;

a pitch detector operatively coupled with the filter, the pitch detector calculating the pitch of the set of residue segments;

an estimator operatively coupled with the pitch detector, the estimator producing a new set of residue segments based upon the set of residue segments and the calculated pitch; and

an inverse filter operatively coupled with the estimator, the inverse filter adding the formant of the consecutive set of audio segments to the new set of residue segments to produce an output audio segment.

22. The apparatus as defined by claim 21 further comprising:

an analyzer operatively coupled with the input, the analyzer calculating formant values for generating the filter.

23. The apparatus as defined by claim 21 further wherein the given audio segment is missing from the plurality of audio segments.

24. The apparatus as defined by claim 21 wherein the audio signal is a voice signal transmitted across a packet based network, the audio signal being a stream of data packets.

25. The apparatus as defined by claim 21 wherein the filter utilizes linear predictive coding filtering techniques.

26. The apparatus as defined by claim 21 wherein inverse filter utilizes linear predictive coding filtering techniques.

27. The apparatus as defined by claim 21 wherein the formant is a function having a variable value across the set of audio segments.

28. The apparatus as defined by claim 21 further comprising:

an overlap add module that applies overlap-add operations to the output audio segment to produce an overlap audio segment.

29. The apparatus as defined by claim 28 further comprising:

a scaler operatively coupled with the overlap add module, the scaler scaling the overlap audio segment to produce a scaled audio segment, the scaled audio segment being the new audio segment.

30. The apparatus as defined by claim 21 further comprising:

an adder that adds the output audio segment to the audio signal in place of the given audio segment.

31. The apparatus as defined by claim 21 wherein the set of consecutive audio segments immediately precede the given audio segment.